

Remarks

The Present Invention and the Pending Claims

The present invention is directed to an improved method and system for disambiguating speech input using multimodal interfaces.

Claims 1-14 are currently pending. Reconsideration and allowance of the pending claims is respectfully requested.

Summary of the Office Action

Claims 1, 4, 6-8, 11 and 14 are rejected under 35 U.S.C. 102(b) as being anticipated by Lai et al., Patent #6006183.

Claim 2 is rejected under 35 U.S.C. 103(a) as being unpatentable over Lai.

Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Lai in view of Bennett et al., Patent #6633846.

Claims 5 and 12-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lai in view of Haddock et al., Patent #5265014.

Claims 9-10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lai in view of Bond et al., Patent #6539348.

Amendments To The Claims

Claims 1 and 11 are currently amended.

Claims 2-10 and 12-14 are retained in their original form.

Support of amendments to claims 1 and 11 is found in paragraphs [0027], [0028] and [0030] of the specification.

Support for the amendment in claim 1 and 11 “wherein the multimodal interaction allows input in voice and visual modes” are found at paragraphs [0027] and [0028].

The office action states: **“Claims 1, 4, 6-8, 11 and 14 are rejected under 35 U.S.C. 102(b) as being anticipated by Lai et al.”**

MPEP section 2131 provides, in pertinent part: “To anticipate a claim, the reference must teach every element of the claim. ...A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference..... The identical invention must be shown in as complete detail as is contained in the ...claim”.

In response to the rejection of claims 1 and 11, claims 1 and 11 have been amended. Applicant’s invention discloses a system for disambiguating speech input using multimodal interaction with an application for disambiguating speech, that is, use of voice and visual input and output modes of user interaction with the multimodal disambiguation mechanism (paragraphs [0023], [0028] and [0030]). In contrast, Lai discloses the use of only a voice input and a speech recognition system wherein text is displayed with attributes (e.g., color) reflecting a recognition confidence score. Therefore, Lai does not disclose, either expressly or inherently the following limitation in claim 1 and 11:

“...*multimodal* interaction,.... wherein the multimodal interaction allows input and output in voice and visual modes;”

Also, in response to the rejection of claims 1 and 11, in applicant’s invention, the output generator generates and presents alternative words or tokens in one or more modes to the user (paragraph [27]) from which the user can select an alternative word, for example a alternative word with the highest confidence level. This functionality is clearly distinguishable from Lai that visually indicates the only the confidence level of speech recognition (column 3, lines 49-55, Lai). For example, see column 4, line 18-23 of Lai that reads: “the graphical user interface application has as its output each word with an associated color and/or attribute. This information is then used to display the word with the associated color and/or attribute, which, in turn, is an indication of the confidence

level associated with each word”. Therefore, Lai does not disclose, either expressly or inherently the following limitation in claim 1 and 11:

“...*multimodal* interaction to present the alternatives to the user and to receive a selection of alternatives from the user ..”.

Furthermore, in response to the rejection of claims 1 and 11, in applicant’s invention, a speech recognition component identifies one or more tokens from user’s speech input, and assigns confidence values to the tokens. The selection component then identifies which two or more of the tokens are to be presented to the user as alternative tokens. The user disambiguation method (paragraph [0027] and [0028]) then interacts with the user by presenting recognition alternative words or tokens to the user for disambiguating speech, and receives from the user, a selection of words or tokens from among the alternative words or tokens as input to the input handler. In contrast, Lai’s speech recognition system merely indicates the confidence level it has in the speech recognition (through colors and/or attributes associated with words), and does not present recognition alternative words or token from which the user can select an alternative word(s) or token. In contrast, the section cited by the examiner (column 3, lines 45-50, Lai) describes a Confidence Level Indicator Process that assigns a color and/or attribute to each word using information from the user control (if any), but does not disclose any selection algorithm that generates recognition alternatives words or tokens. Further, column 3, lines 25-30 in Lai reads “The recognition function translates the acoustic signal into text, i.e., one or more words ... Each word is assigned a confidence level score by a confidence level scorer”. Therefore, in contrast to applicant’s invention which generates alternative words or tokens for each word with a corresponding confidence score, the recognition function in Lai generates text without alternative words or tokens. Therefore, Lai also does not disclose, either expressly or inherently the following limitation in claim 1:

“a selection component that identifies, according to a selection algorithm, which two or more tokens are to be presented to a user as alternatives, wherein said alternatives are words or tokens;”

For the reasons stated above, applicant respectfully submits that claim 1 and claim 11 are novel over Lai, and that the rejection of claims 1 and 11 be withdrawn.

The arguments presented for the allowance of claims 1 and 11 are equally applicable to overcome the rejection of claims 4 and 6. Therefore, applicant respectfully submits that claims 4 and 6 are novel over Lai, and that the rejection of claims 4 and 6 be withdrawn.

Claim 7 is dependent on claim 1. Since claim 1 is not anticipated by Lai, claim 7 that is dependent on claim 1 is also not anticipated by Lai. Therefore, applicant respectfully submits that claim 7 is novel over Lai, and that the rejection of claim 7 be withdrawn.

Claim 8 is dependent on claim 7. Since claim 7 is not anticipated by Lai, claim 8 that is dependent on claim 7 is also not anticipated by Lai. Therefore, applicant respectfully submits that claim 8 is novel over Lai, and that the rejection of claim 8 be withdrawn.

The arguments presented for claims 1 and 11 are equally applicable to overcome the rejection of claim 14. Therefore, applicant respectfully submits that claim 14 is novel over Lai, and that the rejection of claim 14 be withdrawn.

The office action further states: **“Claim 2 is rejected under U.S.C. 103(a) as being unpatentable over Lai”**.

MPEP section 2142 states: “To establish a prima facie case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the references or to combine the reference teachings. Second there must be a reasonable expectation of success. Finally, the prior art references (or

references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must be found in the prior art, not in applicant's disclosure. *In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991)."

First, in response to the rejection of claim 2 over Lai, it is submitted that there is no teaching, suggestion or motivation either in Lai, or in the knowledge generally available to one of ordinary skill in the art, that a speech recognition system displaying text with attributes representing confidence scores as described in Lai can be modified to arrive at the invention recited in claim 2, namely, the disambiguation components and the application residing on a single computing device. Applicant's invention performs disambiguation of speech input based on a multimodal disambiguation mechanism comprising one or more disambiguation components. In contrast, Lai does not disclose a multimodal disambiguation mechanism for speech recognition. Also, neither Lai nor the knowledge generally available teach both the application requiring speech input and the disambiguation components providing disambiguated speech to the application residing on a single computing device. Applicant respectfully submits that since the differences between the present invention and Lai are substantial, the subject matter of the present invention would not have been obvious to a person of ordinary skill in the art at the time the invention was made.

Second, Lai, or the knowledge generally available in the art does not teach the following limitation in claim 2:

"the disambiguation components and the application reside on a single computing device."

For the reasons stated above, applicant respectfully submits that claim 2 is not obvious over Lai and that the rejection of claim 2 be withdrawn.

The office action further states: "**Claim 3 is rejected under U.S.C. 103(a) as being unpatentable over Lai in view of Bennett et al., Patent #6633846.**"

First, in response to the rejection of claim 3 over Lai, in view of Bennett, it is submitted that there is no teaching, suggestion or motivation in Lai or Bennett, or in the knowledge generally available to one of ordinary skill in the art, that a speech recognition system displaying text with attributes representing confidence scores as described in Lai and a distributed real time speech recognition system as described in Bennett can be combined to arrive at the invention recited in claim 3, namely, “the disambiguation components and application residing on separate computing devices.” presenting applicant’s invention, both the disambiguation components and the application contribute to the multimodal disambiguation mechanism. Lai and Bennett et al. do not expressly, or inherently disclose a multimodal disambiguation mechanism (MDM), or the disambiguation components that are part of the MDM. Hence, the distributed nature of the MDM and its components, and the separation of the application and the disambiguation components is not obvious over the cited references.

Second, Lai or Bennett do not teach or suggest the following limitation in claim 3:

“the disambiguation components and the application reside on separate computing devices.”

For the reasons stated above, applicant respectfully submits that claim 3 is not obvious over cited references, and applicant solicits reconsideration of the rejection and allowance of claim 3.

The office action further states: “**Claims 5 and 12-13 are rejected under U.S.C. 103(a) as being unpatentable over Lai in view of Haddock et al., Patent #5265014.**”

In response to the rejection of claim 5 over Lai in view of Haddock et al., it is submitted that Lai or Haddock do not teach or suggest the following limitation in claim 5:

“present the alternatives to the user in a visual form and allow the user to select from among the **alternatives** using a voice input” for disambiguating speech.

This limitation is clearly distinguishable from the sections quoted by the Examiner in Haddock et al. and Lai wherein a referential input, comprising a user selection of **one or more textual elements in previous responses** displayed on the screen, is used to remove referential ambiguity in a database query session; and wherein the user makes the selection by pointing to the desired elements (also see column 4, line 37-40 in Lai). For the reasons stated above, applicant respectfully submits that claim 5 is not obvious over Lai and Haddock and that the rejection of claim 5 be withdrawn.

In response to the rejection of claims 12 and 13 over Lai, first, in view of Haddock, it is submitted that there is no teaching, suggestion or motivation in Lai or Haddock, or in the knowledge generally available to one of ordinary skill in the art, that the system and method of displaying each word with its associated color and/or attribute representing a confidence level as described in Lai (section cited by Examiner: column 3, line 60-64, Lai) and a system and method for receiving a textual input and a referential input to a computer database as described in Haddock, can be combined to arrive at the invention recited in claims 12 and 13, namely, a method of interacting with a user by a concurrent use of multiple modes and receiving a user selection using a combination of voice and visual-based input for disambiguating speech input through a multimodal disambiguation mechanism.. For the reasons stated above, applicant respectfully submits that claims 12 and 13 are not obvious over cited references, and applicant solicits reconsideration of the rejection and allowance of claims 12 and 13.

Second, even if the teachings of Lai and Haddock et al. are combined, the combination that results will be inoperable for the purpose intended by claims 12 and 13. To enable the method and system recited in claims 12 and 13, the following are required: receiving a speech input from a user, determining whether the speech input is ambiguous; if the speech input is ambiguous, performing a multimodal interaction with the user whereby the user is presented with plural alternatives as words or tokens, from which the user can select an alternative. This multimodal interaction comprises “the concurrent use of said visual mode and said voice mode” as recited in claim 12. The user selected alternative is communicated to the application as input to the application. The user selection is through

a “combination of speech and visual-based input” as recited in claim 13. Lai and Haddock et al. do not disclose a system or method that involves simultaneous use of multiple modes of input and output, i.e., a combination of voice and visual modes of input and output, and therefore combining Lai and Haddock et al. will not render operable the method recited in claims 12 and 13. For the reasons stated above, applicant respectfully submits that even if the teachings of Lai and Haddock et al. are combined, they do not support a reasonable expectation of success as required by MPEP section 2142. Therefore, applicant respectfully submits that claims 12 and 13 are not obvious over the cited references, and applicant solicits reconsideration of the rejection and allowance of claims 12 and 13.

The office action further states: **“Claims 9-10 are rejected under U.S.C. 103(a) as being unpatentable over Lai in view of Bond et al., Patent #6539348.”**

First, in response to the rejection of claim 9 over Lai, in view of Bond, it is submitted that there is no suggestion or motivation in Lai or Bond, or in the knowledge generally available to one of ordinary skill in the art, that a speech recognition system displaying text with attributes representing confidence scores as described in Lai and a natural language sentence parser narrowing down the possible interpretations for the words in a sentence can be combined to arrive at the invention recited in claim 9, namely, the one or more disambiguation components disambiguating the alternatives presented to the user in plural iterative stages. In applicant’s invention, the disambiguation components disambiguate the speech recognition alternatives in plural stages, wherein the number of stages is specified by the user or by the type of application. This is clearly distinguishable from Bond, wherein syntactic identifiers (corresponding to parts of speech and other indicators of word usage) of a word are compared with defined rules to narrow down the possible interpretations of the word.

Second, Lai and Bond do not teach or suggest the following limitation in claim 9:

“.. the one or more disambiguation components disambiguates the alternatives in plural iterative stages”

i.e., the one or more disambiguation components operating iteratively to narrow the alternatives is not found in Bond, et al. Also, claim 9 is dependent on claim 1. Since claim 1 is not obvious over Lai, claim 9 that is dependent on claim 1 is also not obvious over Lai in view of Bond et al. Therefore, applicant respectfully submits that claims 9 is novel over Lai and Bond et al., and that the rejection of claim 9 be withdrawn.

Claim 10 is dependent on claim 9. Since claim 9 is not obvious over Lai in view of Bond et al., claim 10 that is dependant on claim 9 is also not obvious over Lai in view of Bond et al. Therefore, applicant respectfully submits that claims 10 is novel over Lai and Bond et al., and that the rejection of claim 10 be withdrawn.

Conclusion

Applicant respectfully requests that a timely Notice of Allowance be issued in this case. If, in the opinion of Examiner Rider a telephone conference would expedite the prosecution of this application, Examiner Rider is requested to call the undersigned.

Respectfully submitted,

Aug. 10, 07
Date

A. Tankha
Ashok Tankha, Esq.
Attorney For Applicant
Reg. No. 33,802
Phone: 856-266-5145

Correspondence Address
Of Counsel, Lipton, Weinberger & Husick
36 Greenleigh Drive
Sewell, NJ 08080
Fax: 856-374-0246